Minority adolescents in ethnically diverse schools: perceptions of equal treatment buffer threat effects

Article (Accepted Version)

CHILD DEVELOPMENT

Minority adolescents in ethnically diverse schools: Perceptions of equal treatment buffer threat effects

Journal:	*Child Development*
Manuscript ID	2015-047.R2
Wiley - Manuscript type:	Special Section
Keywords:	stereotype threat, equal treatment, discrimination
Abstract:	Can perceptions of equal treatment buffer the negative effects of threat on the school success of minority students? Focusing on minority adolescents from Turkish and Moroccan heritage in Belgium (M_age = 14.5; N = 735 in 47 ethnically diverse schools), multilevel mediated moderation analyses showed: (1) Perceived discrimination at school predicted lower test performance; (2) Experimentally-manipulated stereotype threat decreased performance (mediated by increased disengagement); (3) Perceived equal treatment at school predicted higher performance (mediated by decreased disengagement); and (4) Personal and peer perceptions of equal treatment buffered negative effects of discrimination and stereotype threat. Thus, (situational) stereotype threat and perceived discrimination at school both undermine minority student success, whereas perceived equal treatment can provide a buffer against such threats.

Abstract

Can perceptions of equal treatment buffer the negative effects of threat on the school success of minority students? Focusing on minority adolescents from Turkish and Moroccan heritage in Belgium (*M*_age = 14.5; *N* = 735 in 47 ethnically diverse schools), multilevel mediated moderation analyses showed: (1) Perceived discrimination at school predicted lower test performance; (2) Experimentally-manipulated stereotype threat decreased performance (mediated by increased disengagement); (3) Perceived equal treatment at school predicted higher performance (mediated by decreased disengagement); and (4) Personal and peer perceptions of equal treatment buffered negative effects of discrimination and stereotype threat. Thus, (situational) stereotype threat and perceived discrimination at school both undermine minority student success, whereas perceived equal treatment can provide a buffer against such threats.

Minority adolescents in ethnically diverse schools: Perceptions of equal treatment buffer threat effects

In today's increasingly ethnically and racially diverse schools, immigrant children from an early age are vulnerable to social exclusion and discrimination based on minority group membership (Killen, Mulvey, & Hitti, 2013; Killen & Rutland, 2011). Experiences of social exclusion in adolescence cast a shadow over future developmental outcomes, as they typically predict poor academic engagement and performance (Buhs, Ladd, & Herald, 2006; Killen, Rutland, Abrams, Mulvey, & Hitti, 2013; Wentzel, 2009). Following up on earlier research on intergroup exclusion (Killen, Mulvey, et al., 2013), our research examines the intergroup experiences of Turkish and Moroccan minority adolescents, most of whom are Muslims; and the consequences of these experiences for academic engagement and performance. Focusing on adolescence is important because this is the age when young people's views on society and diversity are formed, as they become increasingly aware of their social standing (Rubin, Bukowski, & Parker, 2006), and hence become more susceptible to identity threat.

Identity threat is a key risk factor in adolescence. From a social-identity approach to social development (Killen & Rutland, 2011), minority adolescents will experience identity threat whenever they are faced with discriminatory treatment or negative stereotypes about their group's competence; the latter is commonly termed stereotype threat (Ellemers, Spears, & Doosje, 2002; Steele, Spencer, & Aronson, 2002). Thus, minority children will feel threatened and may disengage from academic work whenever their minority (ethnic, racial or religious) identity is devalued or rejected in their school environment (Baysu, Phalet, & Brown, 2011; Ellemers et al., 2002; Coll et al., 1996). While perceived discrimination is a chronic source of identity threat for members of devalued minority groups (Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak, 2002), stereotype threat is often more situational (Steele,

1997). Both reliably predict minority academic engagement and performance (Steele et al., 2002). Turkish and Moroccan minorities in European educational systems perform worse than their native peers, even when controlling for parental and individual background, and this disadvantage has long-term consequences for unequal chances later in life (Heath & Brinbaum, 2014). Against this background, our first objective was to establish the detrimental effects of perceived discriminatory treatment and situationally-induced stereotype threat on the academic outcomes of Turkish and Moroccan minority adolescents.

In contrast, perceptions of being treated equally communicate that one's group identity is valued (Purdie-Vaughns, Steele, Davies, Ditlmann, & Randall-Crosby, 2008; Tyler, Degoey, & Smith, 1996). Perceived equal treatment in this study refers to the general school climate, how the rules are applied in school, and whether everyone is treated equally or fairly. Our second objective was thus to contextualize identity threat by testing whether a generalized school climate of equal or fair treatment across minority and majority students can counteract this threat. In addition to personal perceptions of equal treatment, we also examined peer perceptions of equality by majority students in the same schools, as a more external contextual measure of intergroup climate at school. We hypothesized that both personal and peer perceptions of equal treatment at school would facilitate minority academic engagement and, subsequently, performance. Moreover, perceived equality was expected to function as a buffer against the negative effects of both chronic threat (perceived discrimination) and situational threat (stereotype threat).

Academic disengagement, the reduction or the absence of effort or persistence in the face of difficulty or failure (Buhs et al., 2006), has been found to predict performance in multiple studies (Fredricks, Blumenfeld, & Paris, 2004). It has also been found to mediate the links between supportive relationships at school and student performance (e.g., Zimmer-Gembeck, Chipuer, Hanisch, Creed, & McGregor, 2006) and between discrimination and

stereotype threat and performance (Keller, 2002; Steele et al., 2002). Accordingly, we expected that both types of threat would harm academic performance via increasing disengagement; whereas perceived equal treatment would enable performance via decreasing disengagement. Overall, this research adds to our understanding of the effects of social exclusion with an experimental demonstration of stereotype threat effects on minority academic outcomes in a real-life setting, and by showing how personal and peer perceptions of equality can counteract those threat effects.

Discrimination and Stereotype Threat

In adolescence, social identity issues become more prominent (Hitti, Mulvey, Rutland, Adams, & Killen, 2014; Horn, 2003), thus making identity threat a key risk factor in this developmental period. Adolescents can experience identity threat when their in-group is devalued in a particular intergroup context, such as at school. This is because people often derive self-worth from their membership in and identification with various social groups (Tajfel & Turner, 1986). Accordingly, when a school explicitly or implicitly communicates disregard or disrespect of minority social identities, it represents an identity-threatening environment for minority students (Derks, van Laar, & Ellemers, 2007; Purdie-Vaughns et al., 2008).

As experiences of discrimination communicate the devaluation of one's social identity, they are seen to pose a chronic threat to the identity of minority group members (Derks et al., 2007). Relatedly, Mendoza-Denton and his colleagues (2002) have argued that one's past experiences of rejection based on one's membership in a devalued group can induce so-called rejection sensitivity, a state of anticipatory threat leading one to 'anxiously expect, readily perceive, and intensely react to status-based rejection' (p. 897). In a longitudinal study among African-American students, they showed that past experiences of racial discrimination, through communicating rejection and inducing feelings of threat,

interfered with school success. Similarly, Benner and Kim (2009) demonstrated negative longitudinal effects of discrimination on school engagement and grades among Chinese-American students.

Perceived discrimination used in this study refers to generalized feelings of peer victimization and exclusion in schools, such as name calling and bullying, and does not refer to ethnic discrimination per se, that is, discrimination based on one's cultural background. However, also more general experiences of peer victimization and exclusion could create identity threat in minority group members, since minority students understand general victimization experiences as instances of discrimination and have been found to be more likely to attribute racial motives to peers' reasons for exclusion, even when non-race based motives were given for the basis for exclusion (Killen, Henning, Kelly, Crystal & Ruck, 2007; Thijs & Verkuyten, 2008).

Even in the absence of overt ethnic discrimination, academic settings may expose minority adolescents to identity threat whenever situational cues – most often implicitly – convey the message that their group is less valued than the majority group (Derks et al., 2007; Purdie-Vaughns et al., 2008). One such situational cue is the presence of negative stereotypes about the minority group's competence. Stereotype threat, as a situational form of identity threat, refers to the presence of salient negative stereotypes about a minority group's competence in a particular domain, such as academic achievement (Steele, 1997). Much experimental evidence has linked stereotype threat to disengagement and performance decrements in minority students (Steele et al., 2002). While both perceived discrimination and stereotype threat, as distinct forms of identity threat, have been found to have detrimental effects on academic outcomes (e.g., Mendoza-Denton et al., 2002; Steele et al., 2002), no study has investigated their *joint* impact. To this end, we investigated simultaneously the effects of discrimination as a chronic threat and stereotype threat as a more situational threat.

Extensive evidence highlights the key role of sustained school engagement in enabling academic success (Fredricks et al., 2004; Zimmer-Gembeck et al., 2006). Conversely, disengagement was reliably associated with decrements in actual academic performance (Buhs et al., 2006). Thus, minority students' performance may suffer in the face of discrimination and stereotype threat to the extent that they are more likely to disengage from general school and class activities or a specific task in a threatening environment (Buhs et al., 2006; Steele et al., 2002; Zimmer-Gembeck et al., 2006). Focusing on disengagement from academic activities in general, Buhs et al. (2006) showed that early peer exclusion and victimization longitudinally predicted disengagement from class activities, which in turn decreased academic achievement. Along those lines, one of the mechanisms behind the detrimental effects of stereotype threat has been shown to be reduced effort, or disengagement from the specific task at hand (Keller, 2002; Steele et al., 2002; Stone, 2002). In line with task-specific disengagement measures in the stereotype threat literature (Keller, 2002; Steele et al., 2002), disengagement in our study refers to reduced effort during a performance task. We predicted that disengagement from the task would mediate the effects of discrimination and stereotype threat on test performance (H1 and H2, respectively).

Equal Treatment as a Buffer

Adolescence is also the developmental period when multiple facets of an intergroup situation are considered simultaneously, including social identity, and perceptions of fairness and equality (Killen, Mulvey, et al., 2013). According to a developmental intergroup perspective, adolescents not only seek and maintain group identity and affiliation, which is regarded as the social domain, but also use notions of fairness and equal treatment in their judgments, which is regarded as the morality domain (Killen & Rutland, 2011; Killen, Rutland, et al., 2013). Adolescents evaluate situations by considering these two domains simultaneously. As such, moral judgments of fairness and equal treatment might offset the

effects of negative intergroup relations such as exclusion and discrimination (Killen, Mulvey, et al., 2013). For instance, it has been shown for majority group members that fairness and equality considerations motivate adolescents to reject group-based exclusion and to promote inclusion (Rutland, Killen, & Abrams, 2010).

A social identity perspective also points to the importance of perceptions of equal treatment (Purdie-Vaughns et al., 2008). Fairness cues perceived by minorities can increase their engagement and performance by communicating the message that their minority identity is valued (Tyler et al., 1996). For instance, Purdie-Vaughns and colleagues (2008) showed that, in an experimentally manipulated 'high-fairness' condition, African-Americans trusted the setting more than those in the 'low-fairness' condition. Importantly, fairness cues offset the threatening interpretation of high-threat cues and increased trust in the organization despite the presence of threat.

Given that perceived equal treatment should increase engagement (Purdie-Vaughns et al., 2008; Tyler et al., 1996) and that academic engagement should enable academic performance (Fredricks et al., 2004; Zimmer-Gembeck et al., 2006), we reasoned that perceived equal treatment at school should reduce disengagement and, in turn, enable better performance. In other words, we predicted that disengagement from the task would mediate the effect of perceived equal treatment on performance (H3). We combined individual-level personal perceptions of equality (H3a) with school-level aggregated minority and majority peer perceptions of equality (H3b) as measures of the intergroup climate in schools.

Moreover, perceived equal treatment should work as a buffer against negative threat effects. From a developmental intergroup perspective (Killen & Rutland, 2011), adolescents differentiate the domain of morality that includes issues such as fairness and equality from the social domain that includes issues such as group identity and norms; and they apply these distinct issues simultaneously to experiences of exclusion or discrimination. Theoretically

therefore, a minority adolescent can be excluded by others (discrimination), or negatively stereotyped in school (stereotype threat), and at the same time perceive the school system as fair or equal for all. This perception of equal treatment in school could then mitigate the negative effects of threat in terms of disengagement and performance. So, we hypothesized that perceived equal treatment would buffer the negative effects of (1) perceived personal discrimination, and (2) of stereotype threat on academic engagement, which in turn should enable better performance (H4 and H5, mediated moderations).

Given developmental changes in social-cognitive and moral reasoning in adolescence (Rutland & Killen, 2015), we explored age-related differences in the hypothesized buffer effect. We tentatively expected that older adolescents would be more susceptible to discrimination as well as to fairness and equality cues at school.

Present Study

The field experiment was part of a large-scale representative survey of ethnically diverse classrooms in secondary schools in Flanders-Belgium (Children of Immigrants Longitudinal Survey - Belgium, 2013). The minority adolescents were the children of Turkish and Moroccan immigrant workers. In addition to their persistent disadvantage in educational and labor market outcomes, Turkish and Moroccan immigrant populations, most of whom are Muslims, are also targets of widespread public prejudice against Muslims in Belgium and in Europe at large (Heath & Brinbaum, 2014). In Western Europe, a large proportion of immigrants are Muslim, while the great majority coming to the U.S. are Christian; and most of Muslim immigrants in Western Europe have a lower socioeconomic profile than those in North America. As a result, the Muslim populations in West European countries are larger than in Canada and the U.S., making religious differences more salient (Alba & Foner, 2014). Given this context, the boundaries between ethnicity and religion are blurred in the case of Turkish and Moroccan minorities in Western Europe (see Umana-Taylor et al., 2014

regarding the ethnic and racial boundaries and how they are blurred in the US). Levels of discrimination and disparagement of Muslim immigrant populations in Western Europe are roughly comparable to the minority status of America's 'involuntary' or 'disparaged' minorities such as African-Americans (Suarez-Orozco, 1991).

Moreover, in the academic domain, less favorable school outcomes relative to the majority population such as lower levels of school performance, less access to academic and higher education go together with negative stereotypes of Muslim students as 'less successful' and 'less intelligent' (Heath & Brinbaum, 2014; Verkuyten & Kinket, 1999). Persistent educational inequalities have long-lasting consequences for their life chances; they are more often unemployed or economically inactive and residentially segregated in deprived urban areas (Heath & Brinbaum, 2014).

Focusing on these devalued minority groups, the field experiment used a typical stereotype threat paradigm where ethnicity is made salient in the experimental condition by filling out questions about ethnicity immediately preceding the test (Steele et al., 2002). This minimal manipulation of stereotype threat resulted in performance decrements for women and minorities in several experimental studies (Ambady, Shih, Kim, & Pittinsky, 2001; Shih, Pittinsky, & Ambady, 1999; Steele & Aronson, 1995). In the control condition, students first took the test and filled out the same ethnicity questions after the test. The test was an inductive reasoning test (CFIT, Cattell & Cattell, 1961) that measures fluid intelligence, as distinct from knowledge-based intelligence tests. The latter type of tests would be less suited because they would rely more heavily on language mastery and they are more sensitive to cultural biases. Thus, CFIT has been shown to be relatively low on cultural bias (Nenty & Dinero, 1981). Moreover, stereotype threat effects in immigrant populations are stronger on fluid intelligence tests than on knowledge-based tests (Appel, Weber, & Kronberger, 2015). Participants also reported their disengagement from the task (Skinner, Kindermann, & Furrer,

2009). Finally, we added self-reported personal experiences of discrimination (Brondolo et al., 2005), and perceptions of equal treatment at school (Gregory, Cornell, & Fann, 2011).

The study covered an academically very heterogeneous population of Turkish and Moroccan minority students who were attending vocational, technical, or academic types of secondary education. The Belgian educational system has a fairly rigid hierarchical tracking structure, which allocates students to different tracks at the beginning of secondary education. Academic tracks prepare students for higher education such as polytechnics or university, whereas vocational tracks lead directly to the labor market. Technical tracks prepare for tertiary education or for work. Students are allocated to different tracks according to their prior school performance; therefore different tracks roughly correspond to different ability groups and determine final educational levels (Baysu & de Valk, 2012).

Method

Participants

After obtaining ethical clearance from the school principal and parental and teacher consent, all eligible students in 47 randomly selected secondary schools in Flanders-Belgium participated in the field experiment during their class hours. Sampled schools were stratified from low (< 10% minority students) over moderate (10-30% and 30-60%) to high levels of ethnic composition (> 60%) using administrative data on foreign languages spoken at home. Within each school, participants were randomly sampled from the first (31%), second (34%) and third year (35%) of lower secondary education. For the purpose of this study, we used self-reported parentage (i.e., one or both foreign-born parents or grandparents) to select Turkish and Moroccan minority students (N = 735). Their ethnic identification —the extent to which they felt they belonged to the respective group — was very high as well (M = 3.60, SD = .53 on a 4-point scale, with 44% of the participants reporting the highest value, that is, 4). Most of these students were also Muslims (98%). Their ages ranged from 12 to 18 (M =

14.49; *SD* = 1.24), with 96% between 12-16 years. Most minority participants in our study were second-generation (81.5%); that is, born in Belgium with one or both parents or grandparents born abroad. First- and third-generation participants were small in numbers (14.9% and 3.6% respectively). Participants attended mainly vocational tracks, with smaller proportions in more selective technical (17%) or academic tracks (34%). There were slightly more boys (59%) than girls (41%). For the measurement of peer perceptions of equal treatment, in addition to Turkish and Moroccan minority peers, we also selected majority Belgian peers in the same classrooms and schools (N = 1287).

It is important to note that this study exclusively focused on Turkish and Moroccan minorities as the focus of the paper was on the consequences of threat for stigmatized minorities, and these are the most stigmatized groups in West European societies (Heath & Brinbaum, 2014). However, as the study was a field experiment conducted in schools, it also sampled other less stigmatized minority students, for instance from EU countries (N = 1135). They were not included in the analyses, because theoretically the stereotype threat condition should not have any significant effects on academic outcomes either for the majority group or for less stigmatized minorities. To show this was the case, separate analyses for majorities and other minorities are available in online supplemental material. Neither showed any significant effects of the stereotype threat manipulation on task disengagement or on non-verbal performance.

Procedure and design

Students completed the test and the questionnaire (in Dutch, the language of schooling in Flanders-Belgium) in class in the presence of a research assistant and a teacher. The research was introduced as part of the international 'Youth in Europe Study' about the life experiences and opinions of youngsters in different European countries. The 47 schools were randomly assigned to either the stereotype threat condition (272 minority participants in 23

schools) or the control condition (425 minority participants in 24 schools) (38 participants had missing values on the condition). In the control condition, students were immediately given the test upon arrival in the classroom. After completing the test, they were given a questionnaire with ethnicity questions such as language spoken at home, religious practice, their parents' or grandparents' country of birth. In the stereotype threat condition, students were asked to fill out the ethnicity questions before undertaking the test. In both conditions, general instructions emphasized that responses would be anonymized and would not be shared with anyone in school. Experimental effects at the school level were controlled for objective school characteristics (i.e., % minority students, tracks, and year). To de-emphasize the evaluative nature of the test, the purpose of the research was introduced as investigating the strategies used to solve Figure Puzzles. The only difference between experimental and control conditions was the order of the test and the ethnicity questions. The assignment of all students in each school to the same condition facilitated the collective administration of the cognitive tests; and it ensured that students and teachers were blind to the different conditions. To assess non-verbal test performance we used a paper and pencil multiple-choice format of an inductive reasoning test with a 7-minute time constraint. Research assistants and teachers present during the research were all majority Belgians.

Measures

Non-verbal test performance was measured as the average score of 27 items from the inductive reasoning subtest of the Culture Fair Intelligence Test (CFIT; Cattell & Cattell, 1961). The items were coded 0 for *incorrect* and 1 for *correct.*

Task disengagement was measured with 4 items that were adapted from Skinner et al.'s (2009) behavioral disaffection scale: During the task…"I acted as if I was working; I was thinking about other things; I did not really do my best; My thoughts wandered off." Task disengagement was measured immediately following the performance task, because this self-

report measure refers specifically to how students behaved during the task. Responses were measured on a 5-point scale, from (1) *not at all* to (5) *very much*, α = .80. The positively worded affective component of task disengagement scale (Skinner et al., 2009) did not predict performance and thus was not included in the analysis.

Perceived discrimination as a source of individual chronic threat was measured with seven items: One explicit discrimination item, "How often are you being discriminated against, treated unfairly, or with hostility at school?", and six items adapted from the 'exclusion/rejection' dimension of the Perceived Ethnic Discrimination Questionnaire–Community Version (PEDQ-CV) (Brondolo et al., 2005): "How often do you experience that other students: bully you; treat you unfairly or in a hostile way; threaten you or bother you; shut you out; call you names or insult you",. All items were rated on a 4-point scale (1) *never* to (4) *always*. These seven items formed a reliable scale α = .89. Importantly, whether using the single discrimination item, the 6 exclusion items, or the combined 7-item scale, the results were similar (this supplemental material is available on request). The one-item of explicit discrimination was significantly correlated with the other 6 exclusion items (r = .36).

Personal and peer perceptions of equal treatment were measured by two items (Gregory et al., 2011): "In my school:" "Some students are allowed to do more than others" (reversed), "The rules are applied equally to all students." Items were measured on a 5-point scale, from (1) *strongly disagree* to (5) *strongly agree*, r = .56. Notice that this measure refers to the school as a whole, in contrast to the measure of perceived discrimination above which refers to each student's own personal experience of discrimination. At the individual level we included only the perceptions of Turkish and Moroccan minority students. At the school level, we aggregated the perceptions of majority students and Turkish and Moroccan minority students in the same schools.

Age was a continuous variable ranging from 12 to 18, excluding the two age outliers (19 and 22 years, $z > 3.5$).

Control variables. To account for compositional differences between schools, relevant student characteristics were included as control variables. Grade retention was included in the analysis to control for individual differences in school success. It was measured with one item: "Have you ever repeated a class" and answers were dummy-coded (1 = *retained*, 0 = *not retained*). Dummy coding was also used for year of education (1^{st} or 2^{nd} year, with 3rd year as a reference category) and school track (vocational or technical, with academic as a reference category). Because of our sampling design, the percentage of students speaking a foreign language at home —as an indication of ethnic composition of schools — was included in the analyses as a control variable; and it was measured by three dummies: 10-30%, 30-60%, and > 60% minority students with < 10% as a reference category, which makes up in total 4 categories. Parental education as a proxy for SES was dropped from the analysis as it did not have significant effects; it did not account for the hypothesized effects either.

Analysis

To test our hypotheses, mediated moderation analysis was conducted with task disengagement as the mediator and test performance as the dependent measure using Mplus 7 (Muthén & Muthén, 1998-2012). In line with current practice in mediation analysis (Rucker, Preacher, Tormala, & Petty, 2011), we focused on testing and presenting the mediation effects (i.e., the indirect effects on performance via disengagement) rather than the unmediated effects on the dependent variable, that is, performance (See Supplemental Online Material for the results of a model without mediation). Multi-level analysis was required because of the nested data structure of students (individual level) within schools (school level) and by the school-level assignment of students to experimental conditions.

We tested (H1) the indirect effect of perceived discrimination on performance via disengagement, (H2) the indirect effect of stereotype threat condition on performance via disengagement; and (H3) the indirect effects of personal perceptions of equal treatment at the individual level (H3a) and peer perceptions of equal treatment at the school level (H3b) on performance via disengagement. Mediated moderation implies that perceived equality would buffer minority performance through reducing disengagement in response to threat. This would require an indirect effect of the discrimination*equal treatment interaction at the individual level (H4, *mediated moderation*) and an indirect effect of the stereotype threat*equal treatment interaction at the school level (H5, *mediated moderation*). The distinction between a mediated moderation and a moderated mediation could be a matter of theoretical preference in line with the research objectives (Muller, Judd, & Yzerbyt, 2005). As we expected that task disengagement would mediate the two moderation effects, it serves mediated moderation goals.

Finally we also explored whether the buffer effect at the within level was the same across different age groups by testing the two-way interactions and the three-way interaction with age (age*discrimination, age*equal treatment, and age*discrimination*equal treatment).

Results

Table 1 provides the means, standard deviations, and correlations for individual and school level variables. Mediated moderation analysis was conducted with task disengagement as the mediator and test performance as the dependent measure in a stepwise approach. We started from a null model with random intercept only. The residual variances of task disengagement and non-verbal test performance were significant both at the individual level [.907 (.056), p <.001; .038 (.003), p <.001] and school level [.056 (.026), p = .03; .005(.002), p = .002], respectively. Intra-class correlations (ICC) indicate that respectively 6% and 11.5% of the total variance in minority disengagement and performance is found between rather than

within schools. In a second step, the model included only control variables (Deviance (-2LL) = 2347.81; AIC = 2399.81; BIC = 2516.38). In a third step, the main effects only model, we added perceived discrimination, stereotype threat condition, and perceived equal treatment as predictors (Deviance (-2LL) = 2123.49; AIC = 2187.50; BIC = 2329.19). The robust Satorra-Bentler scaled chi-square difference test showed that the model fit improved significantly, $\Delta\chi^2(32) = 163.98$, $p < .001$. In order to show the robustness of the main effects without the interactions, in the following, we also reported the hypothesized main effects (H1, H2, H3a and H3b) from the main effects only model in parentheses.

In a final step, we added interaction effects to test our mediated moderation hypotheses, which improved the model fit, $\Delta\chi^2(38) = 29.70$, $p < .001$. Our final model showed the best fit: Deviance (-2LL) = 2098.54; AIC = 2174.54; BIC = 2342.81. So we now report more detailed results from our final model including interactions as shown in Figure 1 (see Appendix 1 for the full model including control variables).

We expected an indirect effect of discrimination on performance via disengagement (H1) but it was not significant ($Z = -0.41$, $p = .681$; in the main effects only model, $Z = -1.63$, $p = .102$). However, minorities' experiences of discrimination had a significant direct effect on non-verbal test performance (Figure 1) so that when minorities experienced discrimination in school, they performed worse on the test.

In line with Hypothesis 2, the experimental condition for stereotype threat at the school level elicited a significant indirect effect on non-verbal performance via task disengagement ($Z = -2.31$, $p = .021$; in the main effects only model, $Z = -2.24$, $p = .025$). In other words, task disengagement fully mediated the impact of stereotype threat on performance: the stereotype threat condition decreased minorities' performance through increasing their disengagement.

As for minorities' personal perception of equal treatment at the individual level (H3a), we found a significant indirect effect on non-verbal performance via disengagement (Z = 2.40, p = .017; in the main effects only model, Z = 2.27, p = .024). In other words, the more equal treatment in schools, as perceived by minorities themselves, the less minority pupils disengaged from the task, which in turn predicted their increased performance on the non-verbal test. As expected, this main effect was qualified by a significant two-way interaction between personal perceptions of equal treatment and of discrimination on task disengagement (Figure 1). In line with the mediated moderation hypothesis (H4), there was a significant indirect effect of this interaction on performance (Z = 2.06, p = .039)

The indirect effect of peer perceptions of equal treatment at the school level (H3b) on nonverbal performance was in the same direction but became non-significant after adding the interactions (Z = 1.40, p = .161; but in the main effects only model, Z = 2.03, p = .043). This main effect was qualified by a significant interaction between the threat condition and peer perceptions of equal treatment on task disengagement (Figure 1). In line with the mediated moderation hypothesis (H5), there was a significant indirect effect of this interaction on performance (Z = 2.016, p = .044).

For the ease of interpretation, we graphed the interactions (see Figures 2 and 3). As expected, when minority and majority students perceive equal treatment in schools, they are protected from the detrimental effects of discrimination and stereotype threat. When minorities perceived high discrimination or experienced stereotype threat, they showed significantly less disengagement when they perceived the school was fair, as compared to unfair (Wald $\chi^2(1) = 47.351$, $p < .001$; Wald $\chi^2(1) = 12.453$ $p < .001$, respectively). Moreover, when the school was considered unfair, higher levels of perceived discrimination and stereotype threat were associated with significantly higher disengagement compared to lower

levels of discrimination or the control condition (Wald $\chi^2(1) = 7.057$, $p = .008$; Wald $\chi^2(1) = 8.447$ $p = .004$, respectively).

We also explored age as a moderator for our effects, to see if considerations of equality were more influential for older adolescents. There was no main effect for age on either task disengagement or non-verbal performance. However, we found a significant three-way interaction of age, discrimination, and personal perceptions of equal treatment on task disengagement ($B = -0.115$, $SE = 0.027$, $p < .001$). Simple effects revealed that the expected buffering effect of perceived equal treatment (Figure 4) against the effect of discrimination on task disengagement only held for the older (mean + 1SD) adolescents. When older minorities perceived high discrimination, they showed significantly less disengagement when they perceived the school was fair, as compared to unfair, (Wald χ^2 (1) $= 41.90$, $p < .001$). For the younger minorities, we see a similar but not significant trend (Wald $\chi^2(1) = 2.99$, $p = .084$). Moreover, when the school was considered unfair, for older adolescents, higher levels of perceived discrimination vs. lower levels of discrimination were associated with significantly higher disengagement (Wald $\chi^2(1) = 9.199$, $p = .002$), but this was not the case for younger adolescents (Wald $\chi^2(1) = 0.143$, $p = .706$). Similar to the two-way interaction of discrimination and perceived equal treatment, we also found a significant indirect effect of this three-way interaction on performance ($Z = 2.06$, $p = .039$) (mediated moderation).

Discussion

The two main objectives of this study were to show the detrimental effects of discrimination and stereotype threat on minority academic outcomes in a real life school context, and to test perceived equal treatment in schools as a protective factor against negative threat effects. We proposed a model where disengagement from the task would mediate these effects on non-verbal test performance. Situationally-induced stereotype threat was associated with higher disengagement from the task which in turn predicted lower performance, while

discrimination only had a direct effect on performance. Perceived equal treatment was related to lower disengagement from the task which, in turn, predicted increased performance. Importantly, perceived equal treatment worked as a buffer against both types of threat: when minority and majority pupils perceived equal treatment in schools, minority pupils were protected from negative threat effects on engagement, and in turn they performed better on the test. Exploring the age-related differences, we also found that this buffering effect worked mostly among the older adolescents who were also more vulnerable to discrimination.

Focusing on Turkish and Moroccan minorities as devalued Muslim minorities in European societies, our research adds to the literature in several ways. *First* and foremost, it adds to the existing research on social identity threat in minority adolescents. This is the first study to test the additive effects of discrimination and stereotype threat, as distinct forms of identity threat (Whaley, 1998). The study also provides the first experimental evidence of stereotype threat in a large random sample of Muslim minority students in real-life school settings in a European migration context where educational opportunities for ethnic minorities are severely restricted (Heath & Brinbaum, 2014). The inclusion of large numbers of academically "less successful" Turkish and Moroccan minority students in mainly vocational tracks widens the scope of existing stereotype threat research beyond highly selective academic environments studied hitherto (Steele & Aronson, 1995). Our findings support the external validity of generic processes of identity threat as an explanation of persistent low performance levels among minority students.

In line with stereotype threat literature (Keller, 2002; Steele, 1997; Steele et al., 2002), we showed that situationally-induced stereotype threat was associated with higher disengagement, and disengagement from the task mediated its effects on performance. A similar indirect effect of discrimination on test performance failed to reach significance, however. This could be due to our measure of disengagement being specific to the task at

hand, rather than a general measure of disengagement from school or class activities or due to the cross-sectional nature of our study. Negative effects of discrimination experiences on performance via increased disengagement might become clearer longitudinally. For instance, Buhs et al. (2006) showed that disengagement from class activities mediated the effects of early peer exclusion and victimization on later school performance.

The direct effect of discrimination on performance was still significant so that the more minority group members perceived discrimination, the lower they performed in the test. Perceived discrimination used in this study refers to peer victimization in schools such as name calling and bullying. Given that there is little research on the effects of peer victimization among minority students (but, see Thijs & Verkuyten, 2008), our findings add to the literature on peer victimization in schools (Monks, Ortega-Ruiz, & Rodriguez-Hidalgo, 2008). The fact that we did not have a measure of ethnic victimization, that is, victimization based on one's cultural background, could be considered as a limitation. Ethnic and general peer victimization have often been studied separately in the literature and both were shown to have negative consequences for minority outcomes (Monks et al., 2008; Thijs & Verkuyten, 2008; Verkuyten & Thijs, 2006; Wolke, Copeland, Angold, Costello, 2013). Possibly, even more pronounced negative effects would emerge if we were to measure ethnic victimization more directly (Monks et al., 2008). Verkuyten and Thijs (2006) found that for Turkish and Moroccan minorities in the Netherlands, these two types of victimization were strongly associated, but not for majority group members or for less stigmatized minorities. Nonetheless, we recommend future research to assess both ethnic and general peer victimization to test their effects simultaneously.

Secondly, our findings underline the importance of equal treatment for minority academic outcomes and as a buffer against negative effects of both chronic and situational forms of threat. Thus, the results add to the growing literature on equal treatment and fairness

from a combined developmental and social identity perspective. The developmental intergroup perspective (Killen, Mulvey, et al., 2013; Killen & Rutland, 2011; Killen, Rutland, et al., 2013; Rutland et al., 2010) focuses on the interplay between moral judgments of equality and fairness vs. group identity and functioning, which becomes most prominent during adolescence. From this perspective, fairness and equal treatment should work as a buffer against exclusion and discrimination, which are foregrounded by a heightened concern of group identity and functioning in adolescence. The social identity perspective (Purdie-Vaughns et al., 2008; Tyler et al., 1996), on the other hand, suggests that fairness and equal treatment give a message about the group's value, and thus should increase trust and commitment. Recall that Purdie-Vaughns et al. (2008) experimentally showed that fairness cues indeed set off threatening situational cues: fairness increased trust in an organization despite the presence of threatening cues.

In keeping with a combined developmental and social identity perspective, we showed that personal and peer perceptions of equal treatment are associated with lower disengagement, and disengagement from the task fully mediates its effects on performance. We also found support for the buffer hypothesis. First, at the individual level, when minority pupils perceive high personal discrimination but think that the students are treated equally in school, they are more resilient against negative discrimination effects; that is, they are less disengaged and in turn perform better on the test. Both discrimination and equal treatment were here measured as personal perceptions and as such they were weakly negatively related.

Going beyond the individual perceptions, we also found that, at the school level, when both minority and majority pupils think that students are treated equally in school, minority pupils are less vulnerable to situationally-induced stereotype threat; in other words, they are less disengaged and, in turn, perform better on the test. Taking into account the actual perceptions of majority group peers in the same intergroup context provides a proper

contextual measure of the intergroup climate. Therefore, both effects considered together provide a stronger case for the importance of equal treatment – especially when it is a norm represented by peer perceptions – as a buffer against social exclusion.

Additionally, given the age range of the adolescents in this study, we investigated age-related differences in the hypothesized effects. We found that the buffering effect of perceived equal treatment against discrimination holds mainly for the older adolescents. From a developmental perspective, since moral reasoning seems to develop well into late adolescence and adulthood (Rutland & Killen, 2015), as one moves forward through adolescence, considerations regarding fairness and equality become more prominent and influential. Moreover, we found that adolescents who were 15 years or older were also more vulnerable to the negative consequences of discrimination - or peer victimization in general - when they considered the school to be unfair. This could be because older minority adolescents -with increasing salience of ethnic identity- have to come to terms with not only the cultural differences between the minority and majority group but also the lower or disparaged status of their own group (Phinney, 1990; Umana-Taylor et al., 2014). As a corollary, middle to late adolescents might more readily expect ethnic victimization and attribute ethnic motives even to more ambiguous situations. In support of this reasoning, Killen and her colleagues (2007) showed that with increasing age, minority adolescents were expecting more race-based exclusions than non-race based exclusions. Finally, it could also be related to the prominence of peer relations during this period (Rubin et al., 2006). It has been shown that the effects of both positive and negative peer relations on school engagement were more pronounced with increasing age during adolescence (Li, Lynch, Kalvin, Liu, & Lerner, 2011). Overall, these findings, although cross-sectional, hint at potential developmental trends in social and moral reasoning of adolescents but obviously longitudinal studies are required to verify these trends.

A longitudinal design would also be required for the most stringent test of the mediational hypotheses, in which the outcome variable, performance in the test, would be measured at a later time point than the mediator, disengagement from the task. However, the hypothesized mediation is in line with the existing finding that increased disengagement mediates the lagged effect of discrimination on minority performance (Buhs et al., 2006). While disengagement was also found to mediate experimental stereotype threat effects (Keller, 2002), there is no evidence of a lagged experimental effect yet. A longitudinal design would be best suited to test the lagged effects of discrimination–though not necessarily of stereotype threat—and whether equal treatment would buffer these effects in the long run.

Another limitation was the lack of a manipulation check. However, the findings in the supplementary material showed that majority students' performance was not affected by the same experimental manipulation. The performance of other less stigmatized minority students was not affected either. This provides further support for the effectiveness of the stereotype threat manipulation and rules out other potential explanations like 'negative mood' or 'tiredness'. If ethnicity questions had put students in a bad mood, which is unrelated to the stereotype threat, these questions would also have negatively affected the performance of less stereotyped minority and majority students.

By randomly assigning schools rather than individual students to conditions, experimental effects could be confounded with variation due to pre-existing differences among schools. The assignment of all students in each school to the same condition, however, facilitated the collective administration of the cognitive tests in large student samples; and it ensured that students and teachers were blind to the different conditions. There was probably adequate power at the School level (ns = ~23 per condition) and we did statistically control for variations at the school level such as tracks and ethnic school composition, which should capture most school-level variation. Moreover, by conducting the experiment in large

heterogeneous samples across highly diverse school settings, we also greatly increased the study's ecological validity. Nevertheless, future studies could usefully assign students rather than schools to stereotype threat and control conditions. An additional limitation of the study is the use of a two-item composite measure to assess perceived equal treatment. Use of a multidimensional scale that taps fairness separately from equal treatment or the experimental manipulation of fairness cues (Purdie-Vaughns et al., 2008) would provide stronger support for the roles of fairness and equal treatment as protective factors in performance contexts.

Finally, our findings should be interpreted in light of the fact that all research assistants and teachers present during the research were majority Belgians. As most teachers in Flanders-Belgium would be majority members, the study context closely mirrors participants' natural school environment. Hence, we did not expect a further increase in threat due to the majority background of research assistants. If teachers or research assistants would have had a minority background, however, this might have diluted situational threat effects. For instance, Marx and Goff (2005) varied the race of a test administrator and showed that Black participants' test performance was less affected by stereotype threat when the administrator was also Black. This could be another protective factor against identity threat.

To conclude, the findings lay the ground for future research which should replicate the protective role of fairness and equal treatment for minority adolescents in performance contexts, where they may be both situationally and chronically exposed to identity threats. The widespread prevalence of exclusion and discrimination experiences and negative stereotypes for many minority groups in Europe today probably conveys a message of devaluation, thus paving the way for academic disengagement and underachievement. Measures to promote fairness and equal treatment in schools would help to break this recursive cycle of devaluation and underachievement and provide the basis for a more inclusive and egalitarian society.

References

Alba, R & Foner, N. (2014). Comparing immigrant integration in North America and Western Europe: How much do the grand narratives tell us? *International Migration Review*, *48* (s1), S263–S291. DOI: 10.1111/imre.12134

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12,* 385-390. DOI: 10.1111/1467-9280.00371

Appel, M., Weber, S. & Kronberger, N. (2015). The influence of stereotype threat on immigrants: A meta-analytic review. *Frontiers in Psychology,* 6:900.DOI: 10.3389/fpsyg.2015.00900

Baysu, G. & de Valk, H. (2012). Navigating the school system in Sweden, Belgium, Austria and Germany: School segregation and second generation school trajectories. *Ethnicities, 12*, 776-799. DOI: 10.1177/1468796812450857

Baysu, G., Phalet, K., & Brown, R. (2011). 'Dual' Identity as a two edged sword: Identity threat and minority school performance. *Social Psychology Quarterly*, *74*, 121–143. DOI: 10.1177/0190272511407619

Benner, A. D. & Kim S. Y. (2009). Experiences of discrimination among Chinese-American adolescents and the consequences for socioemotional and academic development. *Developmental Psychology 45*, 1682-1694. DOI: 10.1037/a0016119

Brondolo, E., et al. (2005). The perceived ethnic discrimination questionnaire: Development and preliminary validation of a community version. *Journal of Applied Social Psychology, 35*, 335–365. DOI: 10.1111/j.1559-1816.2005.tb02124.x

Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom

engagement and achievement. *Journal of Educational Psychology,* 98, 1-13. DOI: 10.1037/0022-0663.98.1.1

Cattell, R.B., & Cattell, A.K.S. (1961). *Test of "g": Culture fair*. Champaign, IL: Institute for Personality Ability Testing.

Children of Immigrants Longitudinal Study (2013). University of Leuven Belgium. [Data file]

Derks, B., van Laar, C. & Ellemers, N. (2007). The beneficial effects of social identity protection on the performance motivation of members of devalued groups. *Social Issues and Policy Review, 1*, 217–56. DOI: 10.1111/j.1751-2409.2007.00008.x

Ellemers, N., Spears, R. & Doosje, B. (2002). Self and social identity. *Annual Review of Psychology, 53*, 161–86. DOI: 10.1146/annurev.psych.53.100901.135228

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59-109. DOI: 10.3102/00346543074001059

Coll, C. G., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., & Garcia, H. V. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, *67,* 1891–1914. DOI: 10.2307/1131600

Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal*, *48*, 904-934. DOI: 10.3102/0002831211398531

Heath A & Brinbaum, Y. (2014). *Unequal attainments*. Oxford University Press.

Hitti, A., Mulvey, K.L., Rutland, A., Abrams, D., & Killen, M. (2014). When is it okay to exclude a member of the ingroup? Children's and adolescents' social reasoning. *Social Development, 23,* 451-469. DOI: 10.1111/sode.12047

Horn, S. S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology, 39,* 71-84. DOI: 10.1037/0012-1649.39.1.71

Keller, J. (2002). Blatant stereotype threat and women's math performance:Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles, 47,* 193–198. DOI: 10.1023/A:1021003307511

Killen, M., Henning, A., Kelly, M.C., Crystal, D., & Ruck, M. (2007). Evaluations of interracial peer encounters by majority and minority U.S. children and adolescents. *International Journal of Behavioral Development, 31,* 491-500. DOI: 10.1177/0165025407081478

Killen, M., Mulvey, K.L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development, 84,* 772-790. DOI: 10.1111/cdev.12012

Killen, M., & Rutland, A. (2011). *Children and social exclusion: Morality, prejudice, and group identity*. New York: Wiley/Blackwell Publishers.

Killen, M., Rutland, A., Abrams, D., Mulvey. K.L., & Hitti, A. (2013). Development of intra- and intergroup judgments in the context of moral and social-conventional norms. *Child Development, 84,* 1063-1080. DOI: 10.1111/cdev.12011

Li, Y., Lynch, A. D., Kalvin, C., Liu, J., & Lerner, R. M. (2011). Peer relationships as a context for development of school engagement during early adolescence. *International Journal of Behavioral Development,* 35, 329-342. DOI: 10.1177/0165025411402578

Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology,* 44, 645– 657. DOI: 10.1348/014466604X17948

Mendoza-Denton, R., Downey, G., Purdie, V., Davis, A., & Pietrzak, J. (2002). Sensitivity to Status-Based Rejection: Implications for African American Students' College Experience. *Journal of Personality and Social Psychology, 83,* 896–918. DOI: 10.1037/0022-3514.83.4.896

Monks, C.P., Ortega-Ruiz, R., & Rodríguez-Hidalgo, A.J. (2008). Peer-victimization in multi-cultural schools in Spain and England. European *Journal of Developmental Psychology*, 5, 507-535. DOI:10.1016/j.pse.2015.02.002

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*, 852–863. DOI: 10.1037/0022-3514.89.6.852

Muthén, L.K. & Muthén, B.O. (1998-2012). *Mplus user's guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Nenty, J. H. & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell culture fair intelligence test using the Rasch model. *Applied Psychological Measurement*, 5, 355-368. DOI: http://dx.doi.org/10.1177/014662168100500309

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: A review of research. *Psychological Bulletin*, 108, 499–514. DOI: 10.1037/0033-2909.108.3.499

Purdie-Vaughns, V. Steele, C. M., Davies, P. G., Ditlmann, R., & Randall-Crosby, J. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology, 94,* 615–630. DOI: 10.1037/0022-3514.94.4.615

Rubin, K.H., Bukowski, W., & Parker, J. (2006). Peer interactions, relationships, and groups. In N. Eisenberg (Ed), *Handbook of Child Psychology: Social, emotional, and personality development, 6th ed.* (pp. 571-645) New York: Wiley.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, *5*, 359–371. DOI: 10.1111/j.1751-9004.2011.00355.x

Rutland, A., & Killen, M. (2015). A developmental science approach to reducing prejudice and social exclusion: Intergroup processes, social-cognitive development, and moral reasoning. *Social Issues and Policy Review,* 9, 121-154. DOI: 10.1111/sipr.12012

Rutland, A., Killen, M., & Abrams, D. (2010). A new social–cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science, 5,* 280–291. DOI:10.1177/1745691610369468

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10,* 80-83. DOI: 10.1111/1467-9280.00111

Skinner, E. A., Kindermann, T. A., & Furrer, C. (2009). A motivational perspective on engagement and disaffection. *Educational and Psychological Measurement*, *69*, 493-525. DOI: 10.1177/0013164408323233

Steele, C. M. (1997). A threat in the air: How stereotype shape intellectual identity and performance. *American Psychologist, 52*, 613–29. DOI: 10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 69,* 797-811. DOI: 10.1037/0022-3514.69.5.797

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology, 34,* 379-440. DOI:10.1016/S0065-2601(02)80009-0

Stone, J. (2002). Battling doubt by avoiding practice: The effect of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin*, *28,* 1667-1678. DOI: 10.1177/014616702237648

Suarez-Orozco, M. (1991). Migration, minority status, and education: European dilemmas and responses in the 1990s. *Anthropology & Education Quarterly, 22*, 99-120. DOI: 10.1525/aeq.1991.22.2.04x0743o

Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 7–24). Chicago, IL: Nelson-Hall.

Thijs, J. & Verkuyten, M. (2008). Peer victimization and academic achievement in a multi-ethnic sample: The role of perceived academic self-efficacy. *Journal of Educational Psychology*, 100, 754-764. DOI: 10.1037/a0013155

Tyler, T.R., Degoey, P., & Smith, H. (1996). Understanding why the justice of group procedures matters: A test of the psychological dynamics of the group-value model. *Journal of Personality and Social Psychology, 70,* 913-930. DOI: 10.1037/0022-3514.70.5.913

Umana-Taylor, A. J. et al. (2014). Ethnic and racial identity during adolescence and into young adulthood: An integrated conceptualization. *Child Development,* 85, 21-39. DOI: 10.1111/cdev.12196

Verkuyten, M., & Kinket, B. (1999). The relative importance of ethnicity: Ethnic categorization among older children. *International Journal of Psychology, 34,* 107-18. DOI: 10.1080/002075999400005

Verkuyten, M. & Thijs, J. (2006). Ethnic discrimination and global self-worth in early adolescence. *International Journal of Behavioral Development,* 30, 107-116. DOI: 10.1177/0165025406063573

Wentzel, K. (2009). Peers and academic functioning at school. In K. H. Rubin, W. M., Bukowski, & B. Laursen (Eds.), *Handbook of Peer Interactions, Relationships, and Groups* (pp. 531-547). New York: The Guilford Press.

Whaley, A. L. (1998). Issues of validity in empirical tests of stereotype threat theory. *American Psychologist, 53,* 679-680.

Wolke, D., Copeland, W. E., Angold, A. & Costello, E. J. (2013). Impact of bullying in childhood on adult. Health, wealth, crime, and social outcomes. *Psychological Science, 24*, 1958-1970

Zimmer-Gembeck, M. J., Chipuer, H. M., Hanisch, M., Creed, P. A., & McGregor, L. (2006). Relationships at school and stage-environment fit as resources for adolescent engagement and achievement. *Journal of Adolescence*, *29,* 911-933.

Table 1

Means, Standard Deviations and Correlations for Turkish and Moroccan Participants at the Individual Level and School Level (N = 735)

	INDIVIDUAL LEVEL	**M (SD)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1.	Age	14.49 (1.24)									
2.	Discrimination	1.35 (0.56)	-.093*								
3.	Equal treatment (personal)	3.52 (1.03)	-.110**	-.171***							
4.	Task Disengagement	2.28 (0.98)	.045	.141***	-.177***						
5.	Performance	0.59 (0.21)	.028	-.196***	.142***	-.212***					
6.	Year1	(30.5%)	-.661***	.086*	.102**	-.044	-.088*				
7.	Year2	(34%)	.011	.042	-.051	.056	-.029	-.475***			
8.	Technical track	(17.1%)	.230***	-.086*	-.019	.007	.148***	-.304***	.057		
9.	Vocational track	(48.9%)	.078*	.139***	-.036	0.51	-.317***	-.010	-.052	-.444***	
10.	Grade retention	(69.6%)	.275***	.051	-.045	.004	-.151***	.007	.006	-.026	.206***
	SCHOOL LEVEL	**M (SD)**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	
11.	Stereotype threat	0.39 (0.49)									
12.	Equal treatment (peer)	3.51 (0.19)	-.152***								
13.	Task Disengagement	2.28 (0.33)	.332***	-.404***							
14.	Performance	0.59 (0.08)	-.098**	.332***	-.632***						
15.	Composition (10-30%)	(11.7%)	-.005	.065^	.125**	.303***					
16.	Composition (30-60%)	(31.3%)	-.204***	-.167***	.120***	-.197***	-.246***				
17.	Composition (60-100%)	(51.2%)	.238***	.122***	-.092*	-.048	-.373***	-.691***			
18.	Technical track (range 0-1)	0.17 (0.19)	.142***	-.267***	.007	.071^	.284***	-.179***	-.102**		
19.	Vocational track (range 0-1)	0.49 (0.29)	.149***	-.220***	.395***	-.692***	-.283***	-.069^	.277***	-.184***	

***p<.001, **p<.01, *p<.05, ^p<.08

NOTE: For dummy coded control variables, percentages are presented rather than means and standard deviations.

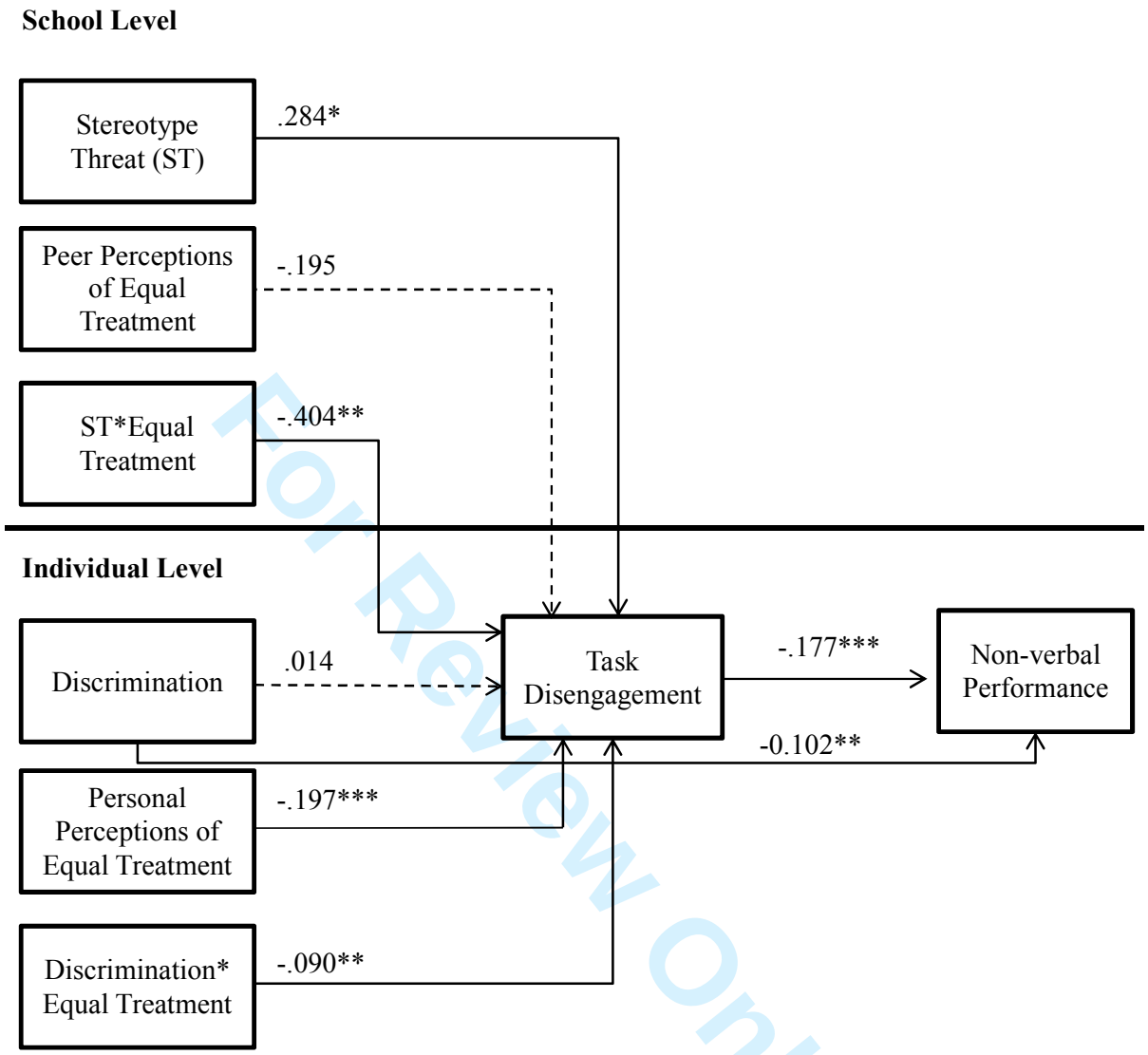

Figure 1. Multilevel Mediated Moderation Model: Task Disengagement as the Mediator and Non-Verbal Performance as the Dependent Variable

Note. The model presents standardized (STDYX) regression results. Age and its interactions were included in the model but are not shown here for simplicity.
***p<.001, **p<.01, *p<.05

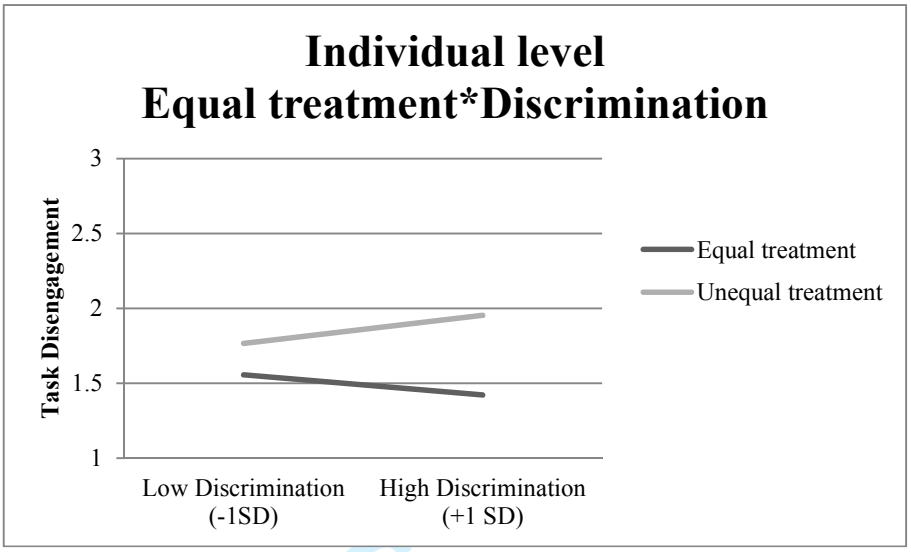

Figure 2. Individual-level buffer effect of equal treatment on the relation between discrimination and task disengagement

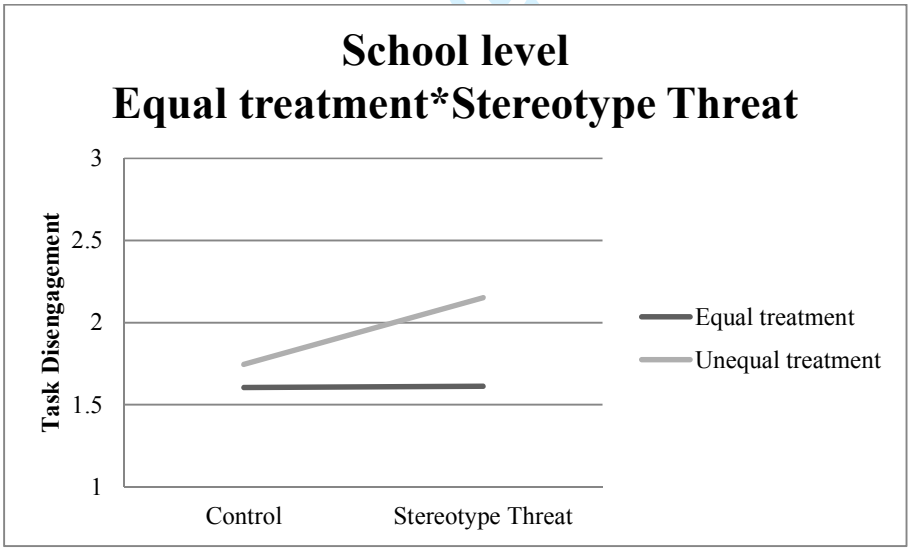

Figure 3. School-level buffer effect of equal treatment on the relation between stereotype threat and task disengagement

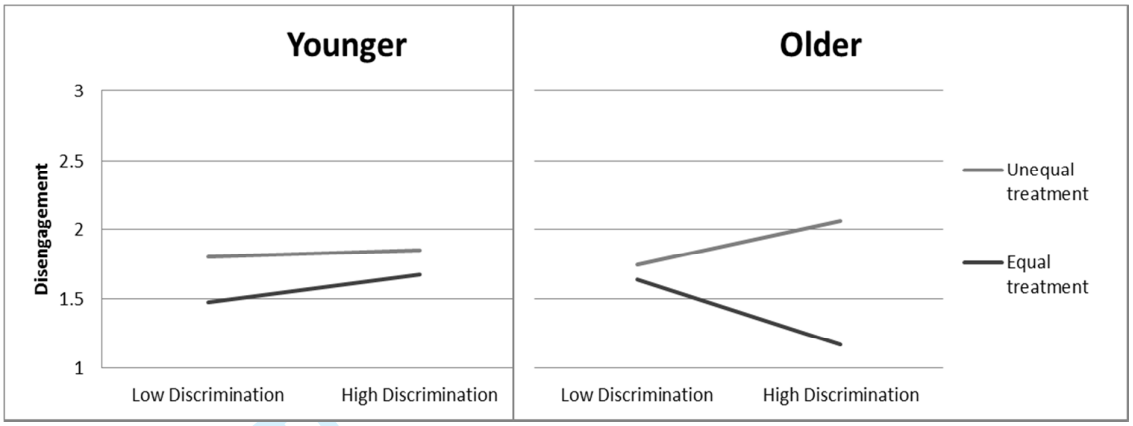

Figure 4. The Individual-level buffer effect of equal treatment on the relation between discrimination and task disengagement for younger and older adolescents

Appendix 1. Multilevel Mediated Moderation Model: Task Disengagement as the Mediator and Non-Verbal Performance as the Dependent Variable

INDIVIDUAL LEVEL	**Mediator Task disengagement**	**Dependent variable Non-verbal performance**
Predictors		
Age	-0.025 (.079)	NS
Discrimination	0.014 (.032)	-0.102 (.036)**
Equal treatment (personal)	-0.197 (.033)***	0.095 (.053)^
Discrimination*Equal treatment	-0.090 (.036)**	NS
Age*Discrimination	-0.051 (.057)	NS
Age*Equal Treatment	-0.065 (.042)	NS
Age*Discrimination*Equal treatment	-0.132 (.027)***	NS
Task disengagement		-0.177 (.054)***
Control variables		
Year1	0.004 (.074)	-0.166 (.036)***
Year2	0.112 (.051)*	-0.116 (.049)*
Technical track	-0.049 (.061)	-0.015 (.040)
Vocational track	-0.062 (.050)	-0.211 (.047)***
Grade retention	NS	-0.076 (.041)^
R-squared	0.082 (.019)***	0.143 (.024)***
SCHOOL LEVEL		
Intercept	4.641 (1.072)***	13.514 (2.142)***
Predictors		
Stereotype threat	0.284 (.139)*	NS
Equal treatment (peer)	-0.195 (.126)**	NS
Stereotype threat *Equal treatment	-0.404 (.164)**	NS
Task Disengagement		-0.692 (.242)**
Control Variables		
Composition (10-30%)	0.706 (0.161)***	0.571 (.283)*
Composition (30-60%)	0.409 (0.223)^	0.159 (.247)
Composition (60-100%)	0.161 (0.190)	0.277 (.242)
Technical track	-0.099 (0.148)	-0.129 (.080)
Vocational track	0.261 (.141)^	-0.540 (.179)**
R-squared	0.928 (.116)***	0.98 (.007)***

Note. Model presents standardized (STDYX) regression results with standard errors in parentheses. NS indicates the 'non-significant' effects that were set to be zero. In this model, unlike the one in Supplemental Material, the effects of the predictors on performance indicate the remaining direct effects in the presence of the mediator.

***p<.001, **p<.01, *p<.05, ^p<.07

Supplemental Online Material (SOM)

Direct Effects Model

In line with current practice in mediation analysis (Rucker, Preacher, Tormala, & Petty, 2011), we focused on testing and presenting the mediation effects (i.e., the indirect effects on performance via disengagement) rather than the unmediated effects on the dependent variable, that is, performance. However for the interested readers, here are the results of the unmediated direct effects model.

Supplemental Material 1. Multilevel Regression Model: Effects on Nonverbal performance

INDIVIDUAL LEVEL	Non-verbal performance
Predictors	
Discrimination	-0.128 (.033)***
Equal treatment (personal)	0.125 (.046)**
Discrimination*Equal treatment	0.012 (.034)
Control variables	
Grade retention	-0.072 (.044)
Year1	-0.164 (.045)***
Year2	-0.123 (.052)*
Technical track	-0.007 (.047)
Vocational track	-0.2 (.049)***
R squared	0.115 (.022)***
SCHOOL LEVEL	
Intercept	10.486 (1.581)***
Predictors	
Stereotype threat	-0.161 (.12)
Equal treatment (peer)	0.086 (.125)
Stereotype threat *Equal treatment	0.383 (.177)*
Control variables	
Segregation (10-30%)	0.102 (.16)
Segregation (30-60%)	-0.101 (.165)
Segregation (60-100%)	0.181 (.142)
Technical track	-0.054 (.123)
Vocational track	-0.716 (.115)***
R squared	0.98 (.006)***

Note. Model presents standardized (STDYX) regression results with standard errors in parentheses. Model fit: AIC = 496.780; BIC = 575.955

***p<.001, **p<.01, *p<.05

Majority Group Sample

Our focus in this paper was the impact of stereotype threat on disadvantaged minorities, thus our analyses were exclusively for Turkish and Moroccan minorities, since theoretically a stereotype threat manipulation should only be relevant for stigmatized minorities, thus should not have any significant effects for majority participants. Nevertheless, since our fieldwork was conducted in schools, majority students (N = 1287) also participated. In this supplemental material, we provided the additional results for majority students. First, majority students' disengagement (M= 2.03, SD = .83) and performance (M = 0.74, SD = 0.15) were significantly different from those of Turkish and Moroccan minority groups (M = 2.28, SD = 1.02; Performance M = 0.59, SD = .21), ps < .001. Secondly, we ran multilevel analyses using Mplus 7 among the majority participants in order to test the effects of stereotype threat condition on their disengagement and performance. Schools were randomly assigned to either the experimental condition (N = 648 majority students in 21 schools) or control condition (N = 590 control in 24 schools). Controlling for grade retention, school year, track, and school level segregation, the experimental condition did not have a significant effect on non-verbal test performance (B = -0.079, SE = .132, p = 0.549) or on task disengagement (B = 0.089, SE = .197, p = 0.650).

Other Minority Group Sample

In addition to the Turkish and Moroccan minorities and majority participants, other minorities also participated in our study (N = 1135). Theoretically, the stereotype threat experimental condition should only have an impact on stigmatized minorities, who feel threatened in response to our manipulation. Therefore, we did not expect the stereotype condition to have any effects on these other non-stigmatized minorities. Moreover, the heterogeneity of the 'other minorities' category makes it harder to interpret any potential

effects of our experimental condition; and other potentially stigmatized groups (e.g. from Africa) are not large enough in size to analyze separately. Nevertheless, in this supplement we provided additional analyses for other minority students. Other minorities (minorities other than Turkish and Moroccan participants) were a heterogeneous group including those from EU and non-EU countries (N = 630 control in 24 schools; N = 475 experimental in 22 schools). First for descriptives, the differences between other minority students' disengagement (M = 2.24; SD=.92) and performance (M = .66, SD = .20) and those of Turkish and Moroccan minorities (M = 2.28; SD= 1.02; M = .59, SD = .21) were only significant for performance, $p < .001$, showing that other minorities significantly performed better than the Turkish and Moroccan minorities. We then ran the multilevel analysis using SPSS 22 among the other minority participants in order to test the effects of stereotype threat condition on their disengagement and performance. Controlling for grade retention, school year, track, and school level segregation, the experimental condition did not have a significant effect on non-verbal test performance (B = -0.003, SE = .015, p = .833) or on task disengagement (B = 0.144, SE = .095, p = .101).